Multiplicative-Binomial Distribution: Some Results on Characterization, Inference and Random Data Generation

Elsayed A.H. Elamir

Department of Mathematics and Statistics, Faculty of Commerce, Benha University, Egypt
&
Management and Marketing Department, College of Business, University of Bahrain,
P.O. Box 32038, Kingdom of Bahrain

shabib40@gmail.com

Multiplicative-binomial distribution is one of the distributions that allows for over-dispersion, and under-dispersion relative to the standard binomial distribution. It will be shown that the multiplicative-binomial distribution can be a very useful model for these situations. Moreover, the confidence interval for the parameters of the multiplicative-binomial distribution is investigated by the profile likelihood methods. The first four moments and simulation procedures for generating data from the multiplicative-binomial distribution using R-software are given. By using four applications to simulated and real data it is shown that the multiplicative-binomial distribution is the same as or outperforming the standard binomial and beta-binomial distributions.

Keywords: binomial distribution; likelihood ratio test; over-dispersion; profile likelihood; zero-inflated data.

1. Introduction

The standard binomial distribution provides a standard framework for the analysis of count data under the assumptions of independent trials and fixed trial probability. In recent years there has been considerable interest in models for count data that allows for over-dispersion, under-dispersion and excess zeros relative to the standard binomial distribution. Application areas are diverse; see, Lambert (1992), Bowman and George (1995), Kibria (2006), Banik and Kibria (2009), Kolev and Paiva (2008) Siegel and Castellan (1988) and Yu and Zelterman (2002, 2008).

The Lovison's multiplicative-binomial distribution (LMBD) is introduced by Lovison (1998) as a generalization of the standard binomial distribution in the case of dependence among the trials that allows for over-dispersion and under-dispersion and includes the standard binomial distribution as a special case. In this paper, it will be argued that the LMBD could be used to model data with over-dispersion and under-dispersion relative to the standard binomial distribution. Moreover, appropriateness of the standard binomial distribution and confidence intervals for the LMBD parameters are investigated by the likelihood ratio test and the profile likelihood method, respectively. The

first four moments and simulation procedures for generating random variates from LMBD using R-software are studied. Four applications to simulated and real data and comparisons with binomial and beta-binomial distributions are given.

In Section 2 the LMBD is reviewed and its moments are obtained. The Likelihood ratio test and profile likelihood for the LMBD are studied in Section 3. The procedures for generating random variates from LMBD are given in Section 4. Four applications to simulated and real data and comparison with binomial and beta-binomial distributions are presented in Section 5. Section 6 is devoted to the conclusion.

2. Lovison's multiplicative-binomial distribution

Consider the random vector $\mathbf{Z} = [Z_1, \ldots, Z_n]$, Z_i being a binary response which measures whether some event of interest is present, '1' or absent, '0' for a sample of n units and let $Y_n = \sum_{i=1}^{n} Z_i$ denote the sample frequency of successes. To accommodate for the possible dependence among trials and under the assumption that the units are exchangeable, Lovison (1998) proposed an alternative form of the multiplicative-binomial distribution introduced by Altham (1978) based on original Cox's representation (1972) as

$$P(Y_n = y) = \frac{\binom{n}{y} \psi^y (1-\psi)^{n-y} \omega^{y(n-y)}}{\sum_{t=0}^{n} \binom{n}{t} \psi^t (1-\psi)^{n-t} \omega^{t(n-t)}}, \quad y = 0, 1, \ldots, n, \tag{2.1}$$

where $0 < \psi < 1$ and $\omega > 0$ are the parameters; for more details about this distribution; see, Lovison (1998). This distribution provides a wider range of distributions than is provided by the standard binomial distribution and includes the standard binomial distribution as a special case at $\omega = 1$.

Figures 1 and 2 show the distribution of Y_n for $n = 10$, $\psi = 0.25$, 0.95, and different values of ω. For small values of ψ and $\omega < 1$ the distribution gives higher probability for zero than the standard binomial distribution. Also, for large values of ψ and $\omega > 1$ the distribution gives higher probability for values at n.

2.1. *Moments of LMBD*

Let

$$n_{(i)} = n(n-1)\cdots(n-i+1), \quad n_{(0)} = 1,$$

$$\pi_i = \psi^i \frac{\kappa_{n-i}(\psi, \omega)}{\kappa_n(\psi, \omega)},$$

and

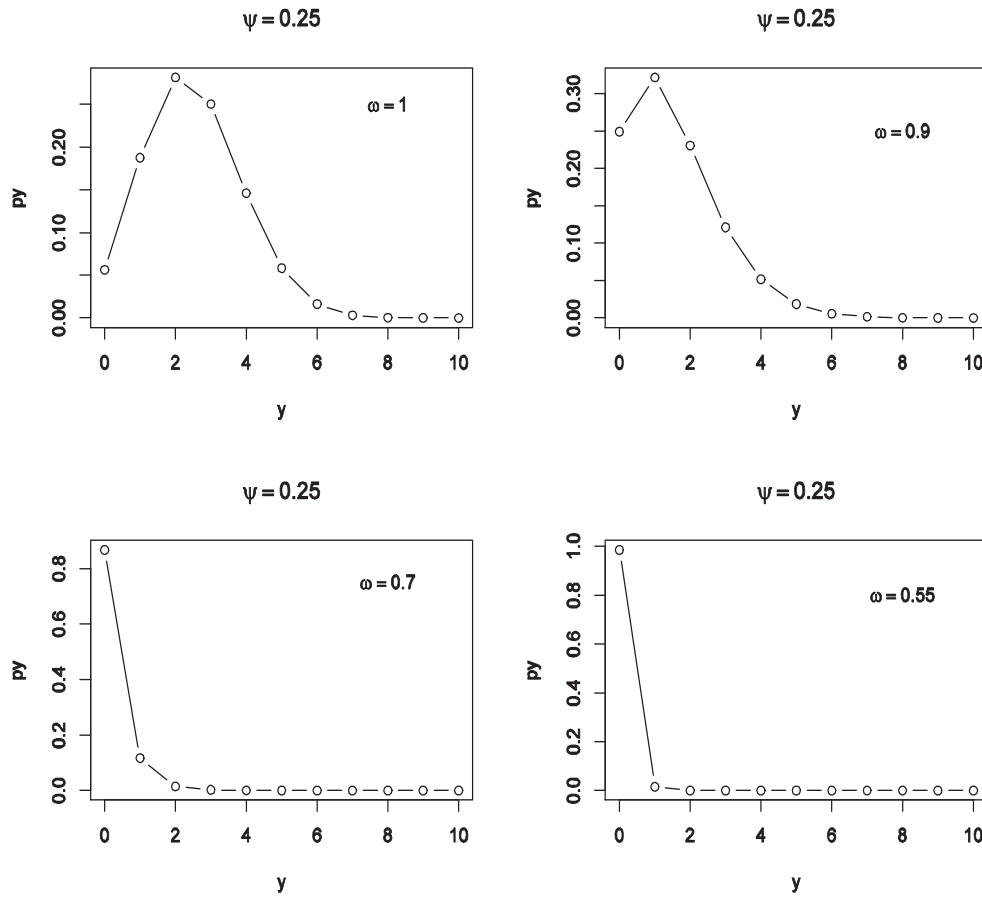$$\kappa_{n-a}(\psi, \omega) = \sum_{x=0}^{n-a} \binom{n-a}{x} \psi^x (1-\psi)^{(n-a-x)} \omega^{(n-a-x)(x+a)}$$

Fig. 1. LMB distribution for selected values of $\omega = 1$, 0.9, 0.7, and 0.56 $\psi = 0.25$ and $n = 10$.

The factorial moments for LMBD can be obtained as

$$\acute{\mu}_{(i)} = E(Y \cdots (Y - i + 1)) = \sum_{i=0}^{n} y \cdots (y - i + 1) \frac{\binom{n}{y} \psi^y (1 - \psi)^{n-y} \omega^{y(n-y)}}{\sum\limits_{t=0}^{n} \binom{n}{t} \psi^t (1 - \psi)^{n-t} \omega^{t(n-t)}} = n_{(i)} \pi_i. \qquad (2.2)$$

For example,

$$\acute{\mu}_{(2)} = E(Y(Y-1)) = \sum_{i=0}^{n} y(y-1) \frac{n(n-1) \binom{n-2}{y-2} \psi^y (1 - \psi)^{n-y} \omega^{y(n-y)}}{y(y-1) \sum\limits_{t=0}^{n} \binom{n}{t} \psi^t (1 - \psi)^{n-t} \omega^{t(n-t)}} = n_{(2)} \pi_2$$

The first four non-central moments are

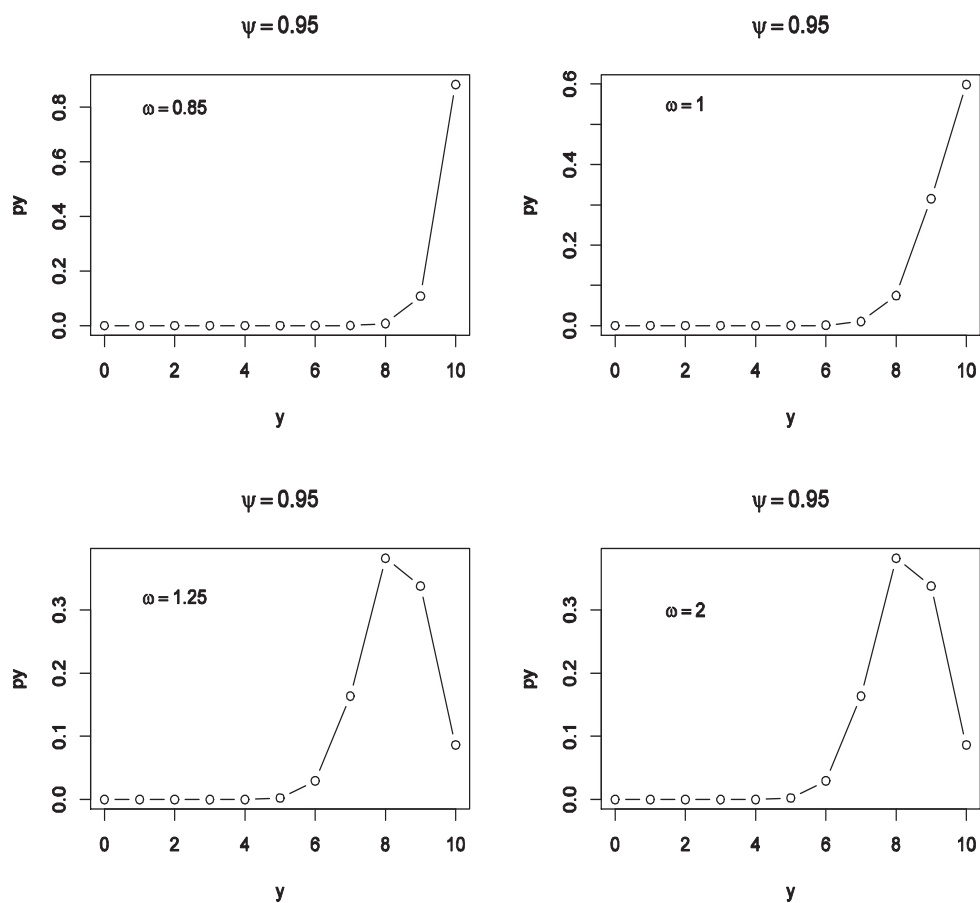$$\acute{\mu}_1 = n_{(1)} \pi_1 = E(Y) = \mu,$$

Fig. 2. LMB distribution for selected values of $\omega = 0.85$, 1, 1.25, and 2, and $\psi = 0.95$ and $n = 10$.

$$\acute{\mu}_2 = n_{(1)}\pi_1 + n_{(2)}\pi_2,$$

$$\acute{\mu}_3 = n_{(1)}\pi_1 + 3n_{(2)}\pi_2 + n_{(3)}\pi_3,$$

and

$$\acute{\mu}_4 = n_{(1)}\pi_1 + 7n_{(2)}\pi_2 + 6n_{(3)}\pi_3 + n_{(4)}\pi_4,$$

The central moments are

$$\mu_2 = n_{(1)}\pi_1 + n_{(2)}\pi_2 - n_{(1)}^2\pi_1^2 = V(Y) = \sigma^2,$$

$$\mu_3 = \left[n_{(1)}\pi_1 + 3n_{(2)}\pi_2 + n_{(3)}\pi_3\right] - 3n_{(1)}\pi_1\left[n_{(1)}\pi_1 + n_{(2)}\pi_2\right] + 2n_{(1)}^3\pi_1^3,$$

and

$$\mu_4 = \left[n_{(1)}\pi_1 + 7n_{(2)}\pi_2 + 6n_{(3)}\pi_3 + n_{(4)}\pi_4\right] - 4n_{(1)}\pi_1\left[n_{(1)}\pi_1 + 3n_{(2)}\pi_2 + n_{(3)}\pi_3\right]$$
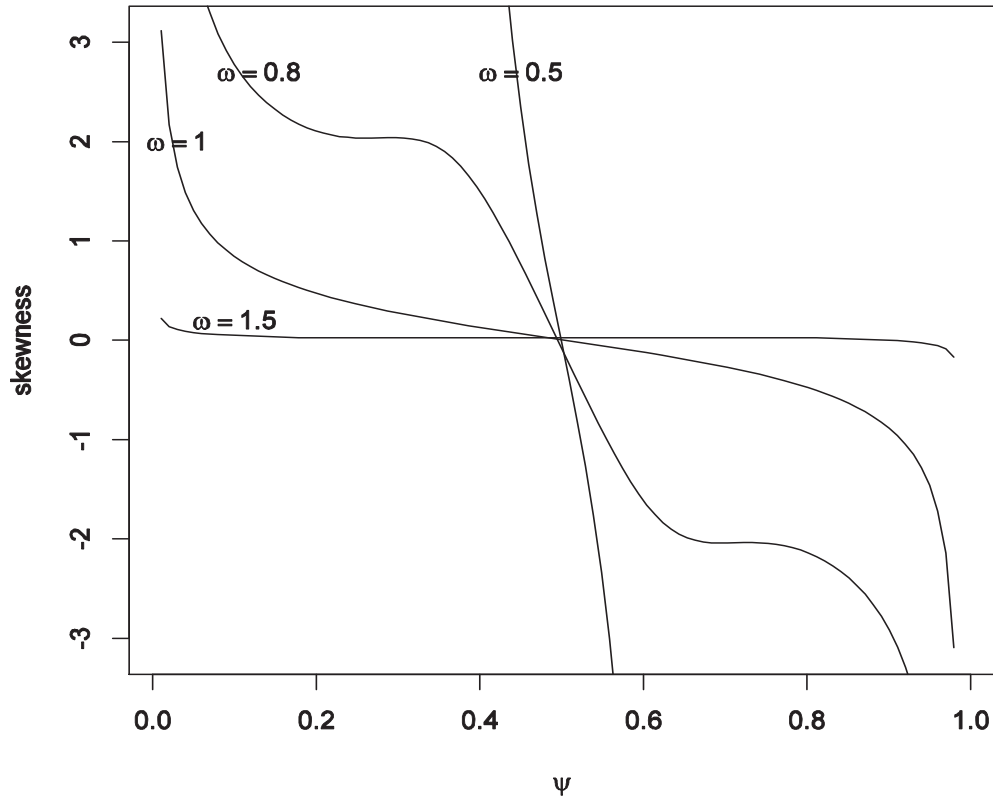$$+ 6n_{(1)}^2\pi_1^2[n_{(1)}\pi_1 + n_{(2)}\pi_2] - 3n_{(1)}^4\pi_1^4$$

Fig. 3. Skewness γ of LMBD with selected values of $\omega = 0.5,\ 0.8,\ 1$, and 1.5 over the range $\psi = (0,1)$ and $n = 10$.

The skewness measure is

$$\gamma = \frac{\mu_3}{(\sigma^2)^{3/2}}.$$

The kurtosis measure is

$$\beta = \frac{\mu_4}{\mu_2^2}.$$

Figures 3 and 4 show the skewness and kurtosis measures for some selected values of ψ and ω at $n = 10$.

From the expected value the relationships among, p, ψ, and ω are

$$p = P(Z = 1) = \psi\,\frac{\kappa_{n-1}(\psi,\omega)}{\kappa_n(\psi,\omega)}, \ \text{ and } \ \psi = p\,\frac{\kappa_n(\psi,\omega)}{\kappa_{n-1}(\psi,\omega)}.$$

Thus p can be considered as a probability of success and ψ as a probability of success weighted by the intra-units association measure ω governing the dependence among the binary responses of
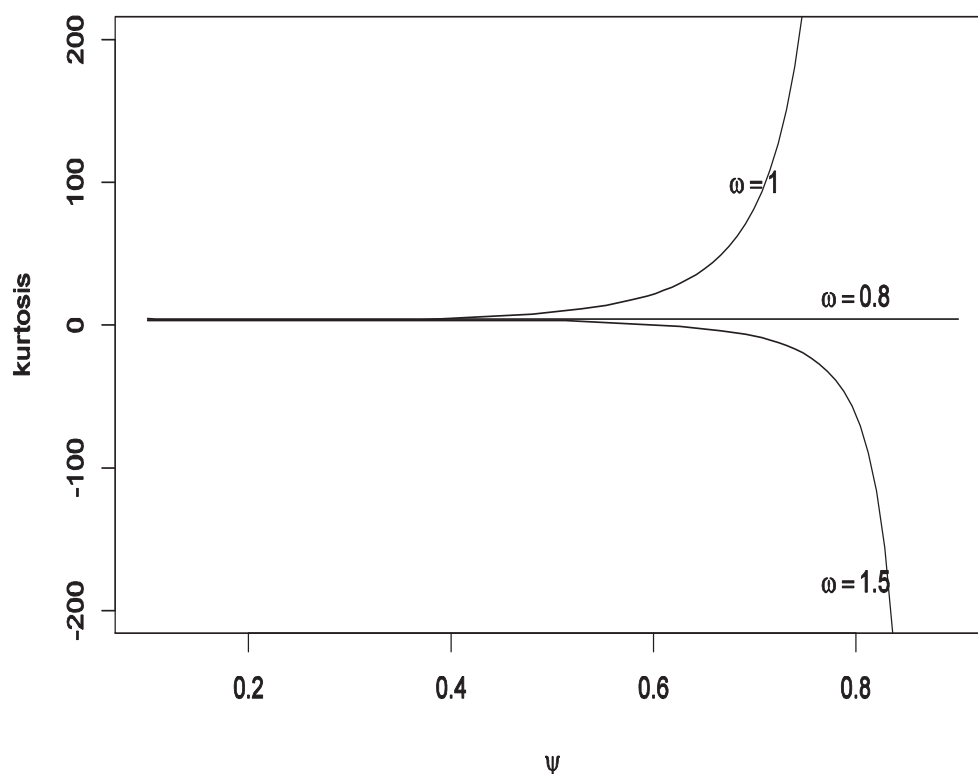
Fig. 4. Kurtosis β of LMBD with selected values of $\omega = 1$, 0.8, and 1.5 over the range $\psi = (0,1)$ and $n = 10$.

the n units. If $\omega = 1$ then $p = \psi$ represents the probability of success of the n units when they are independent. If $\omega \neq 1$ the binary units are not independent and therefore $p \neq \psi$.

3. Likelihood inference on the parameters of the LMBD

When the LMBD fits the data, the profile likelihood and the likelihood ratio can be used to test the appropriateness of the standard binomial distribution and the profile likelihood can be used for estimating the confidence intervals for the LMBD parameters.

3.1. *Confidence intervals based on the profile likelihood*

When the analyst is interested in one of the parameters, the profile likelihood is a useful general tool obtained by maximizing out the remaining parameters for each fixed value of the parameter of interest. It provides a graphical summary of the most plausible values of the parameters of interest and can provide confidence intervals that are more accurate than the standard intervals based on normal theory; see, Pawitan (2001), Severini (2000) and Venzon and Moolgavkar (1988). From

Habib (2010) the logarithm of the likelihood of the multiplicative-binomial can be written as

$$l(\psi, \omega) = \log R! + \sum_{y=0}^{n} f_y \log P(y; \psi, \omega) - \sum_{y=0}^{n} \log f_y!$$

Note that R is the number of sets of n trials, f_y the number of sets giving y success, and $P(y; \psi, \omega)$ is the LMBD. Given the joint likelihood $L(\psi, \omega)$ the profile likelihood of ω can be written as

$$l_p(\omega) = \max_{\psi} l(\psi, \omega), \tag{3.1}$$

where maximization of ψ is performed at a fixed value of ω. To obtain the confidence interval for ω, the profile likelihood is assumed to be an increasing function to the left of its maximum and the following steps are used:

1. Compute the ML estimates $(\widehat{\psi}, \widehat{\omega})$ and the corresponding log-likelihood value,
2. Compute a "reasonable" lower and upper bound ω_1 and ω_2, respectively (e.g., $\widehat{\omega} - 5\, SE(\widehat{\omega})$ and $\widehat{\omega} + 5\, SE(\widehat{\omega})$,
3. Define a grid of values ranging from ω_1 to ω_2.
4. For each grid value ω_i compute the profile log-likelihood value $l_p(\omega_i)$ by maximizing the $l(\psi, \omega_i)$ over ψ-values,
5. Consider the smallest value of ω_i as the lower bound (ω_L) for the $(1 - \alpha)\%$ CI such that $l_p(\omega_i) \geqslant l_p(\widehat{\omega}) - \frac{\chi_{1,1-\alpha}^2}{2} = l(\widehat{\psi}, \widehat{\omega}) - \frac{\chi_{1,1-\alpha}^2}{2}$.
6. Consider the largest value of ω_i as the upper bound (ω_U) for the $(1 - \alpha)\%$ CI such that $l_p(\omega_i) \geqslant l_p(\widehat{\omega}) - \frac{\chi_{1,1-\alpha}^2}{2} = l(\widehat{\psi}, \widehat{\omega}) - \frac{\chi_{1,1-\alpha}^2}{2}$.

If necessary, refine or extend the grid of values around ω_L and ω_U to obtain greater accuracy.

Similarly in which ψ is considered as the parameter of interest and ω is the nuisance parameter, the profile likelihood of ψ can be written as

$$l_p(\psi) = \max_{\omega} l(\psi, \omega). \tag{3.2}$$

To obtain the confidence interval for ψ the following steps can be used:

1. Compute the ML estimates $(\widehat{\psi}, \widehat{\omega})$ and the corresponding log-likelihood value,
2. Compute a "reasonable" lower and upper bound ψ_1 and ψ_2 respectively, (e.g., $\widehat{\psi} - 5\, SE(\widehat{\psi})$ and $\widehat{\psi} + 5\, SE(\widehat{\psi})$,
3. Define a grid of values ranging from ψ_1 to ψ_2.
4. For each grid value ψ_i compute the profile log-likelihood value $l_p(\psi_i)$ by maximizing the $l(\psi_i, \omega)$ over ω-values,
5. Consider the smallest value of ψ_i as the lower bound (ψ_L) for the $(1 - \alpha)\%$ CI such that $l_p(\psi_i) \geqslant l_p(\widehat{\psi}) - \frac{\chi_{1,1-\alpha}^2}{2} = l(\widehat{\psi}, \widehat{\omega}) - \frac{\chi_{1,1-\alpha}^2}{2}$,
6. Consider the largest value of ψ_i as the upper bound (ψ_U) for the $(1 - \alpha)\%$ CI such that $l_p(\psi_L) \geqslant l_p(\widehat{\psi}) - \frac{\chi_{1,1-\alpha}^2}{2} = l(\widehat{\psi}, \widehat{\omega}) - \frac{\chi_{1,1-\alpha}^2}{2}$.

If necessary, refine or extend the grid of values around ψ_L and ψ_U to obtain greater accuracy.

3.2. *Likelihood ratio test for appropriateness of the standard binomial versus LMBD*

Consider a LMB distribution with likelihood function $L(\psi, \omega)$ and $(\widehat{\psi}, \widehat{\omega})$ the maximum likelihood (ML) estimates. The likelihood ratio test statistic (Λ) of the hypothesis $H_0 : \omega = 1$ against $H_1 : \omega \neq 1$ is

$$\Lambda = 2\left[\ln L(\widehat{\psi}, \widehat{\omega}) - \ln L(\ddot{\psi}, 1)\right] = 2\left[l(\widehat{\psi}, \widehat{\omega}) - l(\ddot{\psi}, 1)\right] \tag{3.3}$$

$\ddot{\psi}$ is the ML estimate of the reduced model. From (3.1) the model under the alternative hypothesis is

$$l(\widehat{\psi}, \widehat{\omega}) = \log R! + \sum_{y=0}^{n} f_y \log P(y; \widehat{\psi}, \widehat{\omega}) - \sum_{y=0}^{n} \log f_y!.$$

The model under the null hypothesis is

$$l(\ddot{\psi}, 1) = \log R! + \sum_{y=0}^{n} f_y \log P(y; \ddot{\psi}, 1) - \sum_{y=0}^{n} \log f_y!.$$

Hence,

$$l(\widehat{\psi}, \widehat{\omega}) - l(\ddot{\psi}, 1) = \sum_{y=0}^{n} f_y \log P(y; \widehat{\psi}, \widehat{\omega}) - \sum_{y=0}^{n} f_y \log P(y; \ddot{\psi}, 1).$$

Therefore,

$$\Lambda = 2\sum_{y=0}^{n} f_y \left[\log P(y; \widehat{\psi}, \widehat{\omega}) - \log P(y; \ddot{\psi}, 1)\right].$$

Under the null hypothesis, the test statistic Λ approximately follows a chi-square distribution with one degree of freedom; see, Severini (2000).

4. Generating random data from LMBD using R

The probability mass function for the LMB distribution is

$$P(y; \psi, \omega) = \frac{\binom{n}{y} \psi^y (1-\psi)^{n-y} \omega^{y(n-y)}}{\sum_{t=0}^{n} \binom{n}{t} \psi^t (1-\psi)^{n-t} \omega^{t(n-t)}}, \quad y = 0, \ldots, n.$$

Generating random data from this distribution for given n, ψ, and ω can easily be accomplished in R by using the probability mass function and the sample command, which has the following general syntax.

```
>\#x:a vector of elements from which to choose, size: \# of item to choose
>sample(x, size, replace = FALSE, prob = NULL)
```

Example 4.1. This example is to draw 10000 and 100000 sample data from LMB distribution with $n = 4$, $\psi = 0.20$, and $\omega = 0.45$. From the probability mass function of LMBD we obtain the probability distribution in Table 1.

The following command is used.

Table 1. Exact and simulated mean and variance from LMBD with $n = 4$, $\psi = 0.20$, and $\omega = 0.45$.

y	0	1	2	3	4
$P(y)$	0.89597	0.08164	0.01377	0.00510	0.00349
		Simulated Size			
	Exact	10000	100000		
Mean	**0.13851**	0.13780	0.13977		
Variance	**0.21949**	0.22303	0.21941		
Skewness	**4.47101**	4.50310	4.47905		

```
>x=sample(c(0,1,2,3,4),size,replace=TRUE,prob=
c(0.89597,0.08164,0.01378,0.00510,0.00350))
```

The size can be replaced by the desired value of 10000 and 100000 and therefore the variable *x* contains random data drawn from the LMBD with $n = 4$, $\psi = 0.20$, and $\omega = 0.45$. The mean, the variance and skewness of simulated data are compared with exact distribution in Table 1. A program in R-software to generate random data from LMBD (lmbd.sim) is given as follows:

```
>MBD.dist=function(n,y,p,w) \{\#prob. of LMB distribution

\#n:\#of trials, y:random variable, p:psi value, w:omega value\#

t=0:n;
t1=choose(n,y)*p^y*(1-p)^(n-y)
            *w^(y*(n-y));

t2=choose(n,t)*p^t*(1-p)^(n-t)
            *w^(t*(n-t));
t1/sum(t2) \}

>lmbd.sim=function(r,n,p,w)\{\#Simulation data from LMBD

\#r:\# of observations, n:size, p:psi value, w:omega value

yr=0:n; pr=MBD.dist(n,yr,p,w); sample(yr,r,replace=TRUE,prob=pr)\}
```

5. Applications to simulated and real data

Four applications to simulated and real data are studied. The first application deals with data simulated from a standard binomial. The second application presents data coming from a distribution which has variance less than the standard binomial. The third and fourth applications deal with real data with excess-zeros, which usually have more variance than the standard binomial. Moreover the useful and common-used beta-binomial distribution that is defined as
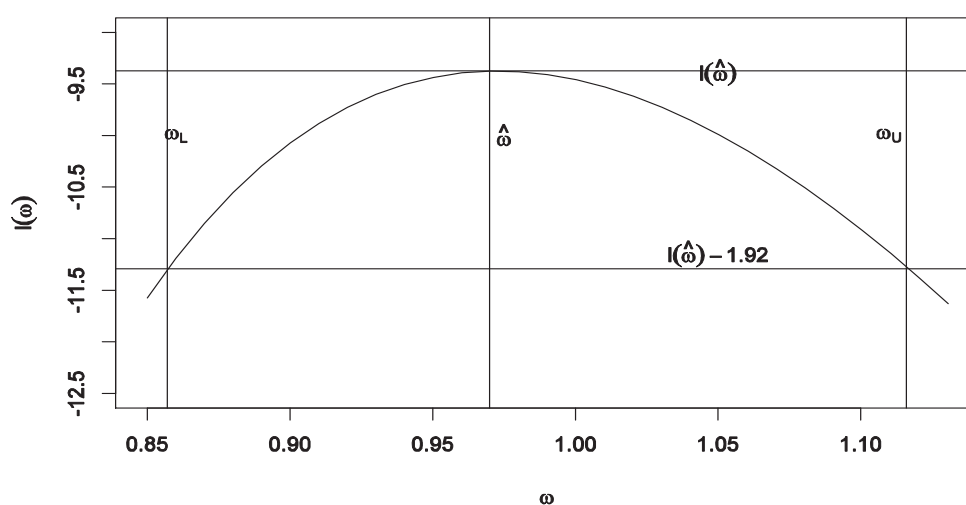
$$f(x; \alpha, \beta) = \binom{n}{x} \frac{B(\alpha + x, \beta + n - x)}{B(\alpha, \beta)}, \quad x = 0, 1, \ldots, n,$$

is compared with LMBD in both simulated and real data analysis. The package "VGAM" in R-software is used to estimate the parameters α and β.

Table 2. Simulated data from binomial distribution with $n = 4$ and $p = 0.15$.

y	0	1	2	3	4	Total
Obs.	255	186	52	6	1	500
E. LMBD	256.46	183.67	52.42	{7.07	0.38}	$X^2 = 0.068$
E. binomial	253.71	187.58	52.01	{6.41	0.30}	$X^2 = 0.032$
E. Beta-binomial	254.87	185.75	52.30	{6.74	0.33}	$X^2 = 0.003$

Obs.: observed and E.: Expected value.

Fig. 5. Profile likelihood and 95% CI: (ω_L, ω_U) for the data in Application 1.

5.1. *Application 1: Simulated data from binomial*

The data are simulated from a standard binomial distribution with $n = 4$ and $p = 0.15$. The expected values and Pearson's X^2 statistic are given in Table 2. From Habib (2010) the maximum likelihood estimates for LMBD parameters are $\widehat{\psi} = 0.164$ and $\widehat{\omega} = 0.97$. The values of X^2 statistic in Table 2 show a suitable fit for the data by all the distributions.

Figure 5 shows the profile log-likelihood function for ω with its maximal value of -9.373, and the 95% confidence interval $(0.857, 1.116)$. Clearly, this C.I. suggests that a standard binomial distribution would be an appropriate model for the data, since $\omega = 1$ is included in the interval. This can also be seen by the likelihood ratio test, $\Lambda = 2(0.0827) = 0.1654$ indicating the appropriateness of the standard binomial model.

5.2. *Application 2: Simulated data from LMBD*

The data are simulated from a LMBD with $n = 6$, $\psi = 0.5$, $\omega = 1.25$, and variance 0.937. The expected values and X^2 statistics are given in Table 3. From Habib (2010) the maximum likelihood estimates for LMBD parameters are $\widehat{\psi} = 0.494$ and $\widehat{\omega} = 1.22$. The estimate for $\widehat{p}$ is 0.496. The values of X^2 statistics in Table 3 show a suitable fit for the data by the LMBD but not by the

Table 3. Simulated data from LMBD with $n = 6$, $\psi = 0.5$, and $\omega = 1.25$.

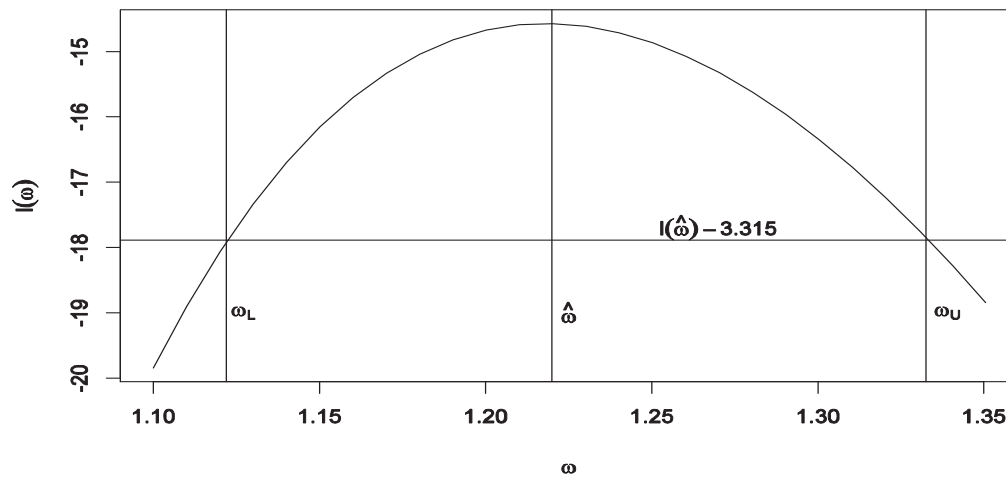| y	0	1	2	3	4	5	6	Total
Obs.	3	25	129	194	122	26	1	500
E. LMBD	{1.81	28.35}	125.05	198.27	119.2	{25.76	1.56}	$X^2 = 0.44$
E. Binomial	8.19	48.39	119.05	156.22	115.30	45.39	7.44	$X^2 = 38.8$
E. Beta-binomial	Not defined							

Fig. 6. Profile likelihood and 99% CI: (ω_L, ω_U) for the data in Application 2.

binomial model and the parameters of beta-binomial distribution cannot be estimated using package "VGAM".

Figure 6 shows the profile log-likelihood function for ω with its maximal value of -14.57, and the 99% confidence interval $(1.122, 1.332)$. Clearly, this C.I. suggests that a standard binomial distribution would not be an appropriate model for the data, since $\omega = 1$ is not included in the interval. This can also be seen by the likelihood ratio test $\Lambda = 2(22.44) = 44.88$.
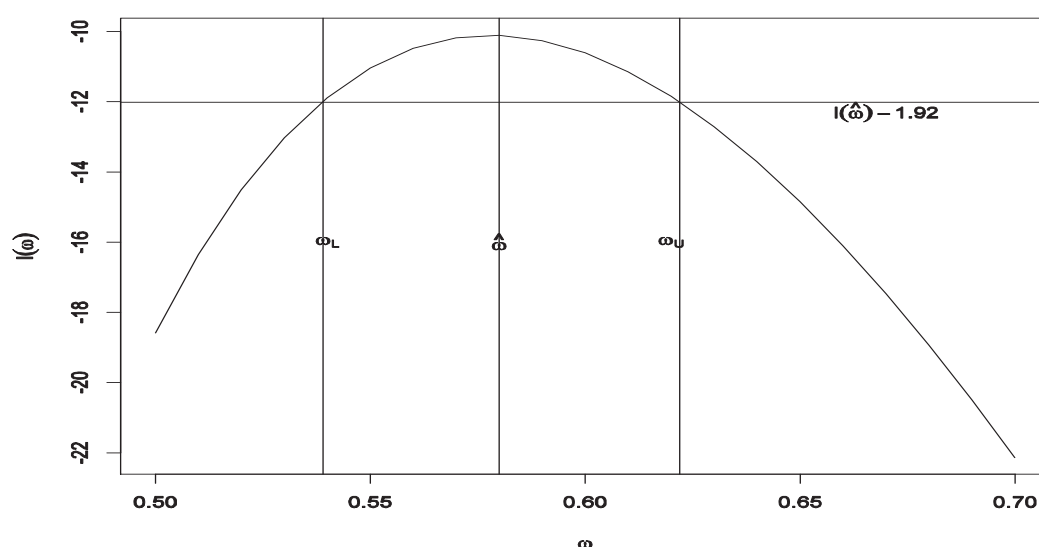
5.3. *Application 3: Traffic accident research*

Kuan *et al.* (1991) discuss data coming from the California Department of Motor Vehicles master driver license file. The variable of interest is the number of accidents per driver. A possible motivation can be seen in the possibility of finding risk factors involved in the accidents. The data, the expected values and X^2 are given in Table 4. These data have mean 0.203 and variance 0.235 and there is a clear spike of extra zeros (about 83%) representing the accident-free drivers. From Habib (2010) the maximum likelihood estimates for LMBD parameters are $\widehat{\psi} = 0.145$ and $\widehat{\omega} = 0.58$. The estimate for $\widehat{p}$ is 0.067. The values of X^2 statistics in Table 4 show a suitable fit for the data by the LMB and beta-binomial distributions but not by the binomial distribution.

Figure 7 shows the profile log-likelihood function for ω with its maximal value of -10.10, and the 95% confidence interval $(0.539, 0.622)$. Clearly, this C.I. suggests that a standard binomial

Table 4. Distribution of the number of accidents per deriver.

number of accidents	0	1	2	3	Total
Number of drivers	4499	766	136	21	5422
Relative frequency	0.830	0.141	0.025	0.004	
E. LMB	4499.38	770.07	130.60	21.95	$X^2 = 0.286$
E. binomial	4403.56	948.68	{68.12	1.63}	$X^2 = 334.9$
E. Beta-binomial	4497.7	764.65	142.81	16.82	$X^2 = 1.36$

Fig. 7. Profile likelihood function and 95% CI: (ω_L, ω_U) for the traffic accident research data.

distribution would not be an appropriate model for these data. This can also be seen by the likelihood ratio test $\Lambda = 2(80.16) = 160.33$.

5.4. *Application 4: Crime Sociology*

In a study on deviating behavior Bohning *et al.* (1997) provide an analysis of a data set on 4301 persons with criminal behavior. The variable of interest is the number of criminal acts per person. The motivation for such a study might be to find factors leading to deviating behavior. The data, the expected values and X^2 statistic are given in Table 5. These data have mean 0.0776 and variance 0.1209. There is a clear spike of extra zeros (about 94%) representing acts-free criminals. From Habib (2010) the maximum likelihood estimates for LMBD parameters are $\widehat{\psi} = 0.203$ and $\widehat{\omega} = 0.460$. The estimate for $\widehat{p}$ is 0.0155. The values of X^2 statistics in Table 5 show a suitable fit for the data by the LMBD but not by the binomial model for all level of significance and the beta-binomial distribution at 10% and 5% level of significance.

Figure 8 shows the profile log-likelihood function for ω with its maximal value of -13.32, and the 95% confidence interval (0.435, 0.485). Again, this C.I. suggests that a standard binomial

Table 5. Distribution of the number of criminal acts.

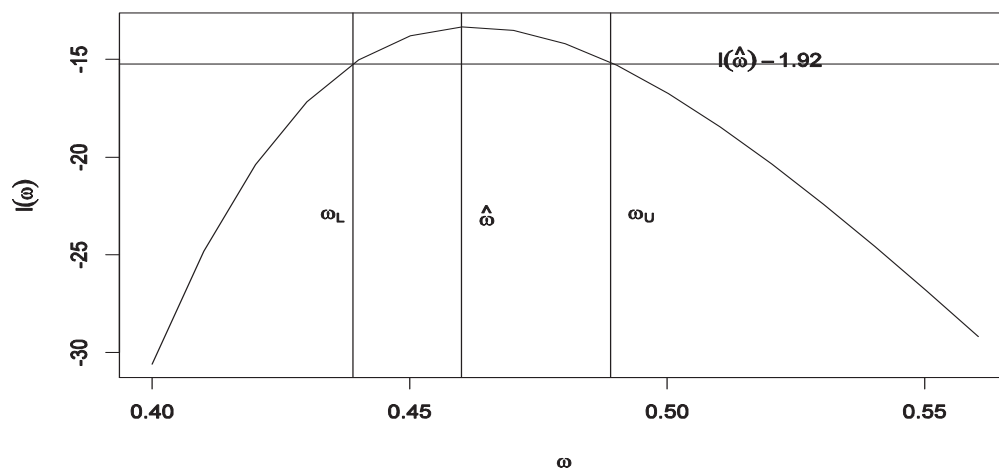| # of criminal acts	0	1	2	3	4	5	Total
# of persons	4037	219	29	9	5	2	4301
Relative frequency	0.9386	0.0509	0.0067	0.0021	0.0012	0.0005	
E. LMBD	4032.66	229.38	24.66	6.27	{3.76	4.27}	$X^2 = 2.55$
E. Binomial	3977.8	313.13	{9.86	0.155	0.0012	0}	$X^2 = 151.3$
E. Beta-binomial	4030.66	218.40	42.14	{8.34	1.33	0.12}	$X^2 = 8.1$

Fig. 8. Profile likelihood function and 95% CI: (ω_L, ω_U) for the crime sociology data.

distribution would not be an appropriate model for these data. This can also be seen by the likelihood ratio test $\Lambda = 2(114.95) = 229.9$.

6. Conclusion

It has been shown that the LMB distribution can be very useful in a variety of applications since it allows for over-dispersion and under-dispersion relative to the standard binomial distribution. The first four non-central and central moments are obtained. Moreover, the LMBD has been used for testing the appropriateness of the standard binomial distribution to the data using the likelihood ratio test and the confidence interval has been obtained for its parameters using the profile likelihood. The simulation procedures for generating data from multiplicative-binomial distribution using R-software are illustrated. It is shown that the LMB distribution could outperform the standard binomial and beta-binomial distributions using four applications to simulated and real data.

Acknowledgment

References

[1] Altham, P. (1978). Two generalizations of the binomial distribution. *Applied Statistics*, **27**, 162–167.

[2] Bohning, D., Dietz, E. and Schlattmann (1997). Zero-inflated count models and their applications in public health and social science. In: *Applications of latent trait and latent class models in the social sciences*, ed. by Jurgen R. and Langeheine, WAXMANN Verlag.

[3] Banik, S. and Kibria, B.M.G. (2009). On some discrete distributions and their applications with real life data. *Journal of Modern Applied Statistical Methods*, **8**, 423–447.

[4] Bowman, D. and George, E.O. (1995). A saturated model for analyzing exchangeable binary data: application to clinical and development toxicity studies. *Journal of American Statistical Association*, **90**, 871–879.

[5] Cox, D.R. (1972). The analysis of multivariate binary data. *Applied Statistics*, **21**, 113–120.

[6] Habib E.A. (2010). Estimation of log-linear binomial distribution with applications. *Journal of Probability and Statistics*, **10**, 1–13.

[7] Kolev, N. and Paiva, D. (2008). Random sums of exchangeable variables and actuarial applications. *Insurance: Mathematics and Economics*, **14**, 147–153.

[8] Kibria, B.M.G. (2006). Applications of some discrete regression models for count data. *Pakistan Journal of Statistics and Operation Research*, **2**, 1–16.

[9] Kuan, J., Peck, R.C. and Janke, M.K. (1991). Statistical methods for traffic accident research, in: *Proceedings of the 1990 Taipei Symposium in Statistics*, June 28-30, 1990, ed. by Min-Te Chao and Philip E. Cheng. Taipei, Institute of Statistical Science.

[10] Lambert, D. (1992). Zero-inflated Poisson regression, with application to defects in manufacturing. *Technometrics*, **34**, 1–14.

[11] Lovison, G. (1998). An alternative representation of Altham's multiplicative-binomial distribution. *Statistics & Probability Letters*, **36**, 415–420.

[12] Pawitan, Y. (2001). *In all likelihood: Statistical modeling and inference using likelihood.* (1st Ed.), Clarendon Press, Oxford.

[13] Severini, T.A. (2000). *Likelihood methods in statistics.* (1st Ed.), Oxford University Press.

[14] Siegel, S., and Castellan, N.J. (988). *Nonparametric statistics for the behavioral sciences* (2nd Ed.), New York, NY: McGraw-Hill.

[15] Venzon, D.J. and Moolgavkar, H. (1988). A Method for computing profile-likelihood-based confidence intervals. *Applied Statistics*, **37**, 87–94.

[16] Yu, C. and Zelterman, D. (2002). Sums of dependent Bernoulli random variables and disease clustering. *Statistics & Probability Letters*, **57**, 363–373.

[17] Yu, C. and Zelterman, D. (2008). Sums of exchangeable Bernoulli random variables for family and litter frequency data. *Computational Statistics & Data Analysis*, **52**, 1636–1649.